



Towards Smart Visualization Framework for Climate Simulations

Lokman Rahmani, Matthieu Dorier, Luc Bougé, Gabriel Antoniu, Robert Sisneros, Tom Peterka

► To cite this version:

Lokman Rahmani, Matthieu Dorier, Luc Bougé, Gabriel Antoniu, Robert Sisneros, et al.. Towards Smart Visualization Framework for Climate Simulations. 2016. hal-01290268

HAL Id: hal-01290268

<https://inria.hal.science/hal-01290268>

Preprint submitted on 17 Mar 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Joint Laboratory
for Petascale Computing



Inria



Towards Smart Visualization Framework for Climate Simulations

Lokman Rahmani, Matthieu Dorier, Luc Bougé
Gabriel Antoniu
Roberto Sisneros
Tom Peterka

ENS Rennes, IRISA
INRIA Rennes
University of Illinois at Urbana Champaign
Argonne National Laboratory

Outline

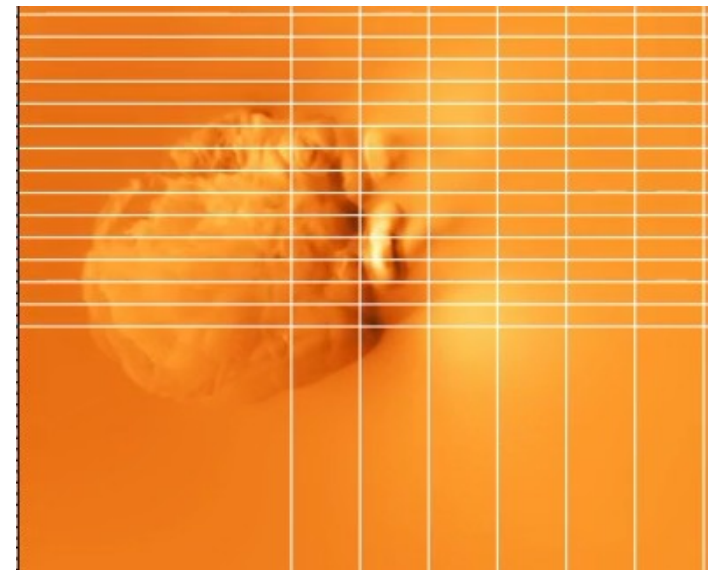
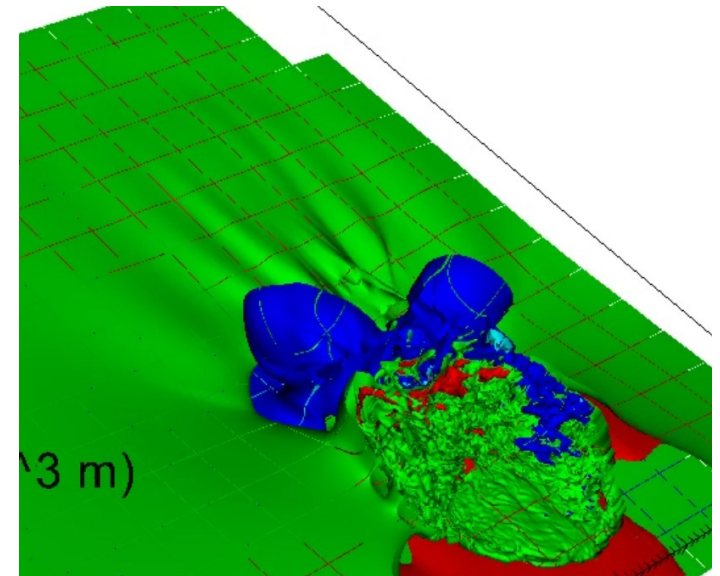
1. HPC Climate Simulations & In Situ Visualization
2. Smart In Situ Visualization
3. Experimental Results
4. Conclusion

Outline

1. HPC Climate Simulations & In Situ Visualization
2. Smart In Situ Visualization
3. Experimental Results
4. Conclusion

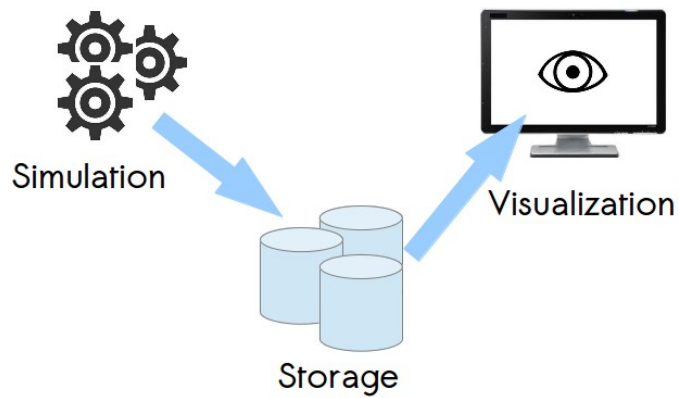
CM1 : Atmospheric Phenomena Simulation

- CM1 on Kraken
 - Simulated space : 48 x 44 x 200 per process
 - 9216 cores
 - 14.2 GBytes/Iteration (~1.69 MBytes for each compute process)
- CM1 on BlueWaters
 - Simulated space : 3840 x 3840 x 400
 - 6400 cores
 - 23.6 GBytes/Iteration (~3.7 MBytes for each process)

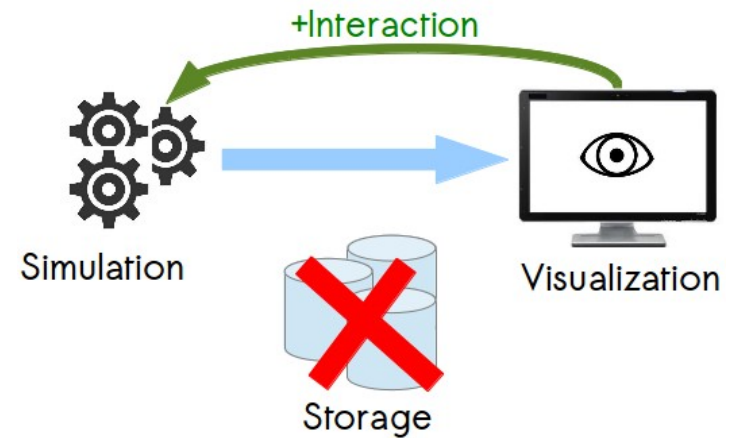


In Situ Visualization

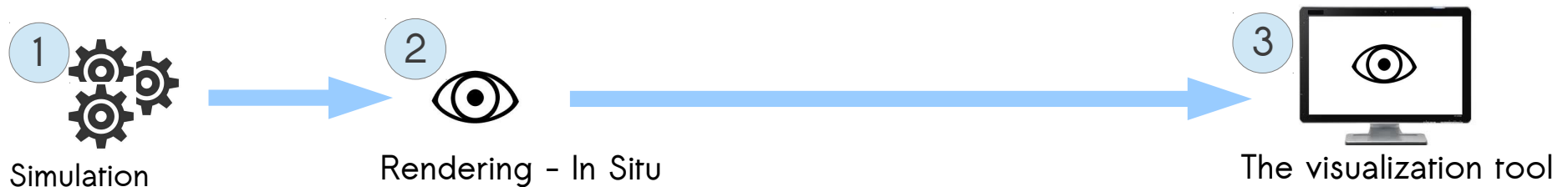
Off-line



In Situ

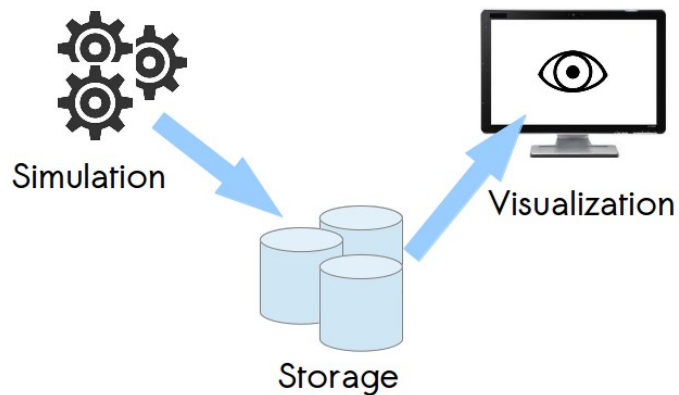


In Situ Visualization Workflow

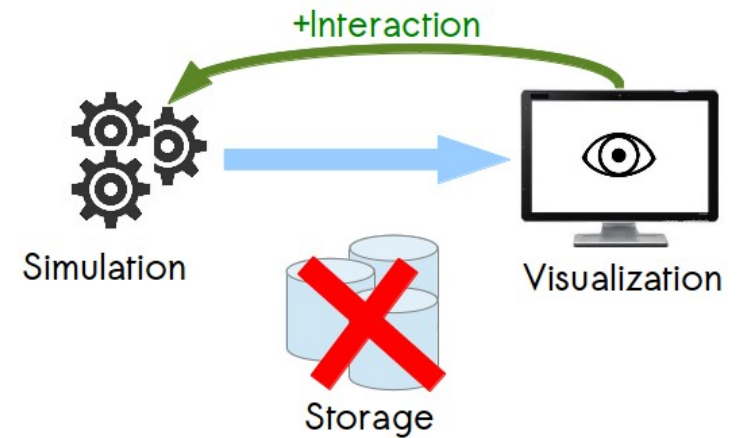


In Situ Visualization

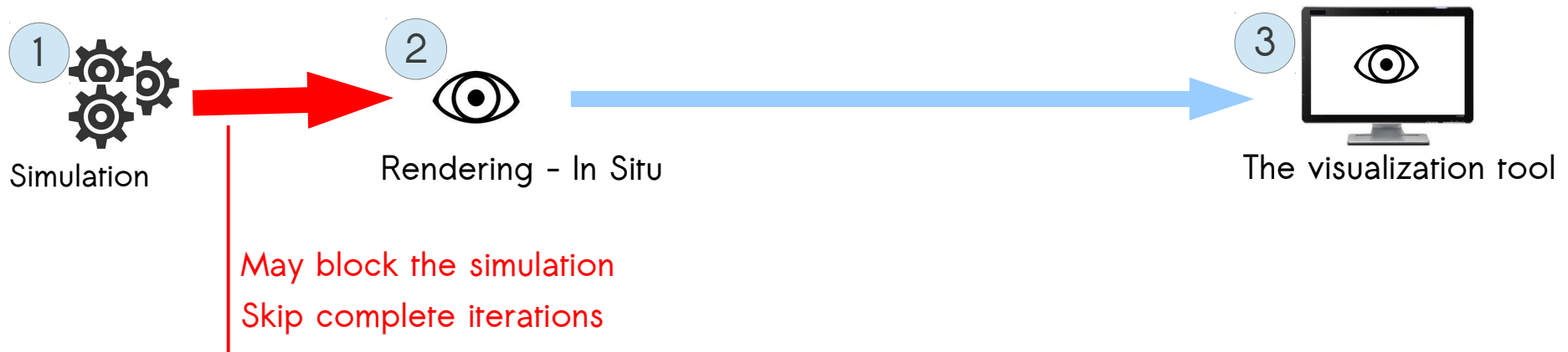
Off-line



In Situ

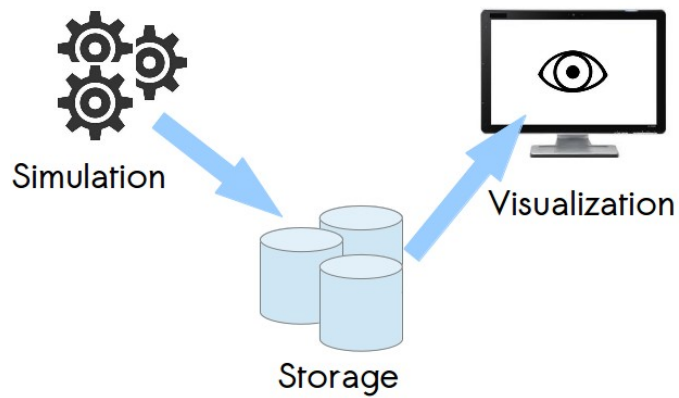


In Situ Visualization Workflow

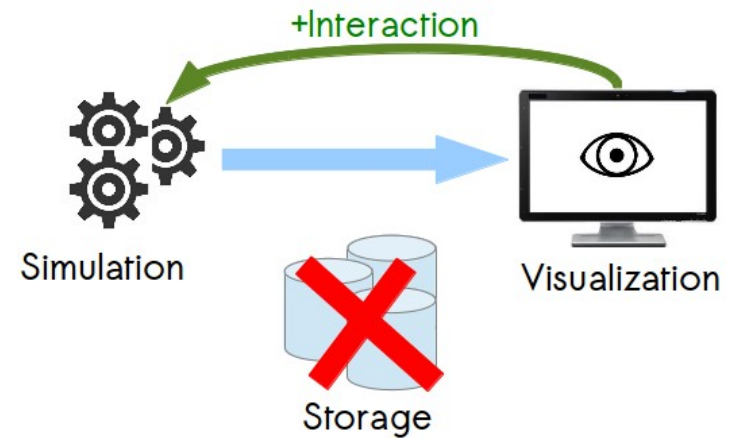


In Situ Visualization

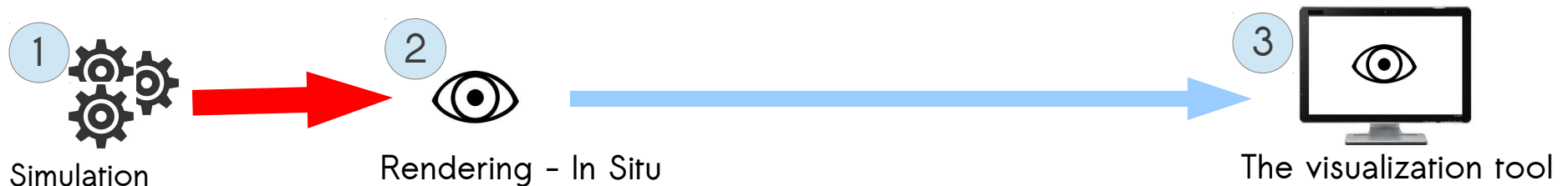
Off-line



In Situ

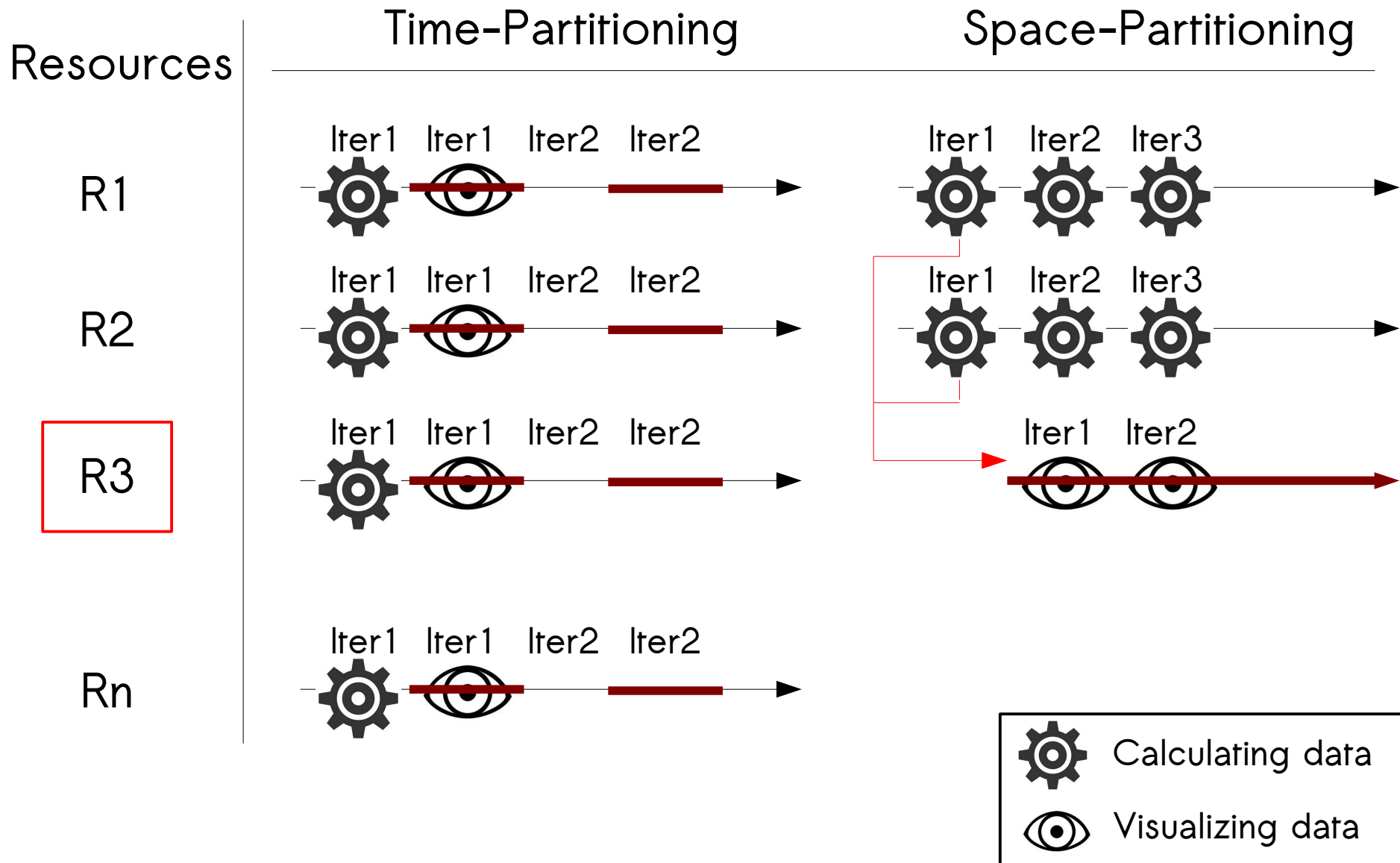


In Situ Visualization Workflow

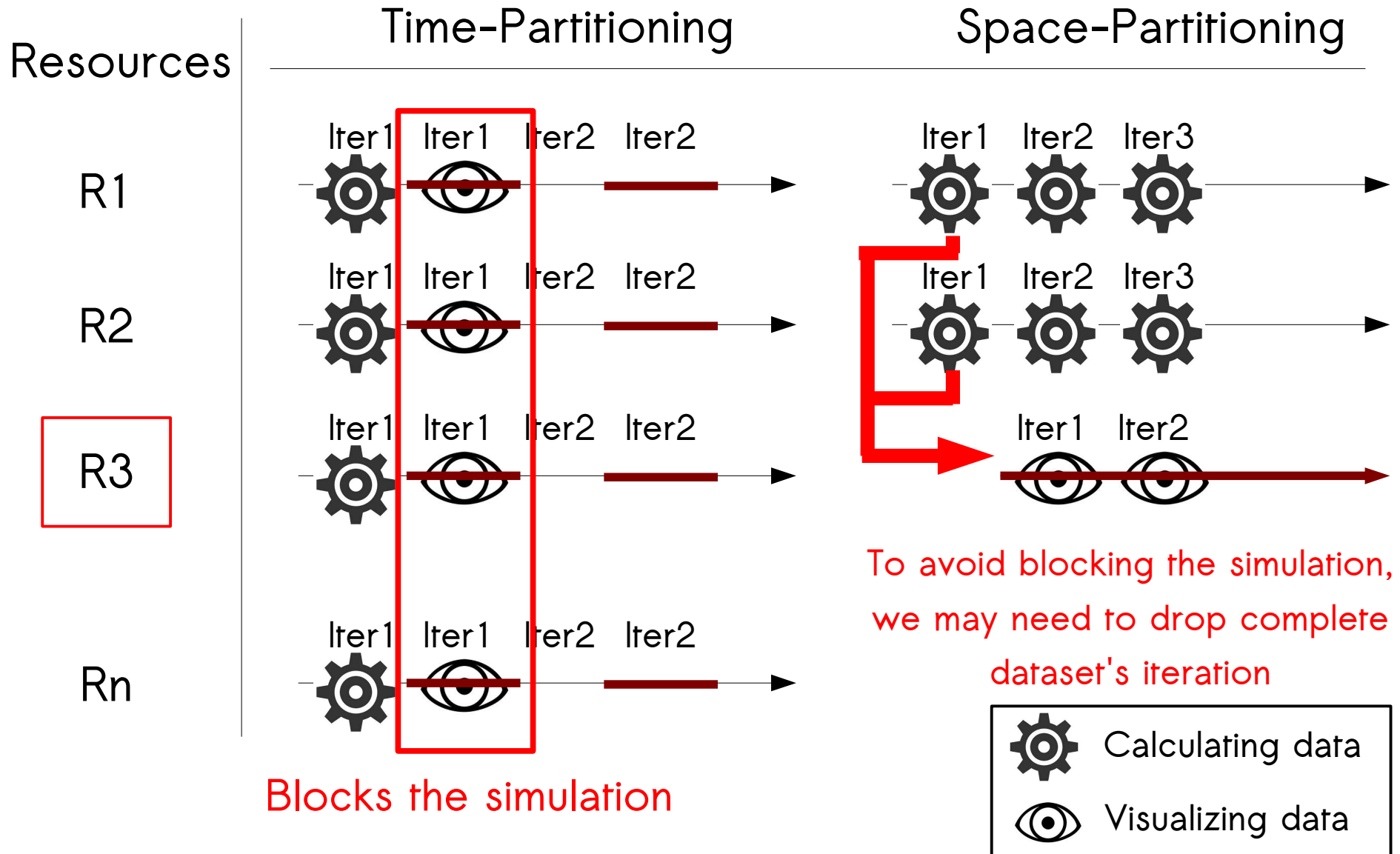


Avoid blind drop of data

In Situ Visualization Implementation



In Situ Visualization Implementation

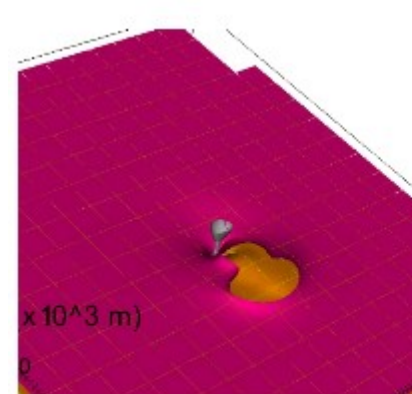
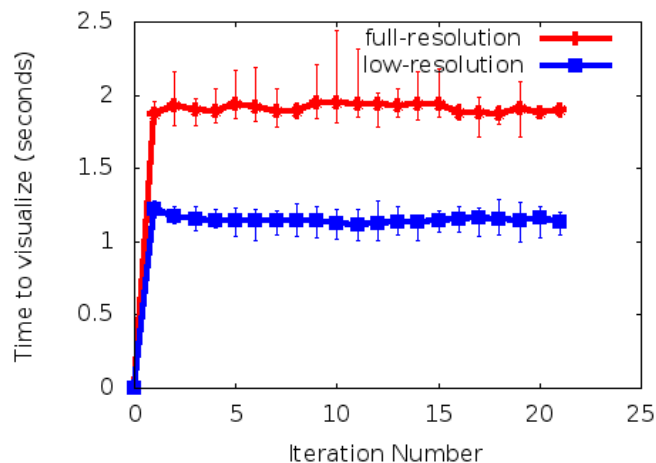


Be Smart, Visualize only relevant data

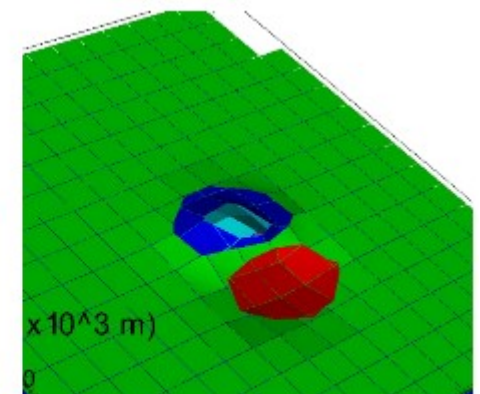
- Avoid blind drop of data
 - Keep all iterations
 - Reduce the amount of data on each iteration
- On each iteration, visualize only relevant data considering the physical phenomena being simulated
 - Interesting data will be visualized in full resolution
 - The rest will be visualized in low resolution

Be Smart, Visualize only relevant data

- Avoid blind drop of data
 - Keep all iterations
 - Reduce the amount of data on each iteration
- On each iteration, visualize only relevant data considering the physical phenomena being simulated
 - Interesting data will be visualized in full resolution
 - The rest will be visualized in low resolution



(a) Full resolution



(b) Low resolution

Be Smart, Visualize only relevant data

- Avoid blind drop of data
 - Keep all iterations
 - Reduce the amount of data on each iteration
- On each iteration, visualize only relevant data considering the physical phenomena being simulated
 - Interesting data will be visualized in full resolution
 - The rest will be visualized in low resolution

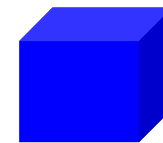
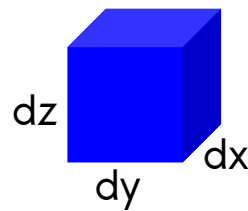
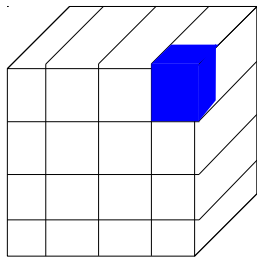
What is 'relevant' data?
and
how to detect it?

Outline

1. HPC Climate Simulations & In Situ Visualization
2. Smart In Situ Visualization
3. Experimental Results
4. Conclusion

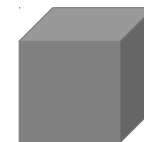
What is Relevant in a Dataset?

- The solution needs to be:
 - Generic: application-independent
 - Automatic: user-transparent
 - Efficient
- Defined semantic : Variation of the data



full resolution:

$$dx \times dy \times dz$$



low resolution:

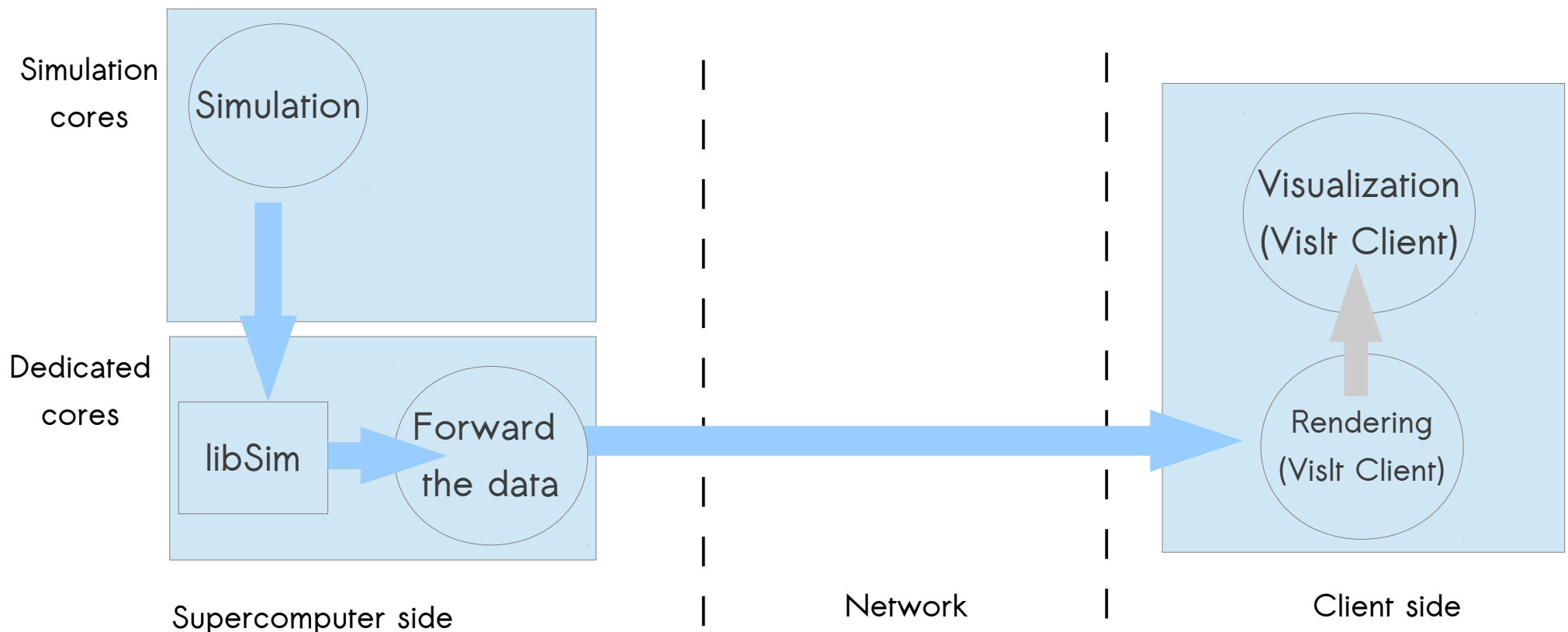
$$2 \times 2 \times 2$$

How to Automatically Detect Relevant Data?

- Metrics used to do filtering should :
 - Provide different levels of variation measurement
 - Can be 'normalized'
- To detect variation in \mathcal{X} , a set of values $\{x_1, x_2, \dots, x_n\}$
 - Entropy (Information Theory) : of a random variable \mathbf{X}
 - Coefficient of Variation (Statistics) $C_v = \frac{\sigma}{\mu}$
 - Independent from the unit of measurement
 - Hard to normalize
 - Gradient (Image Processing)
 - Calculate the derivate over each axis
 - Is not a complete metric

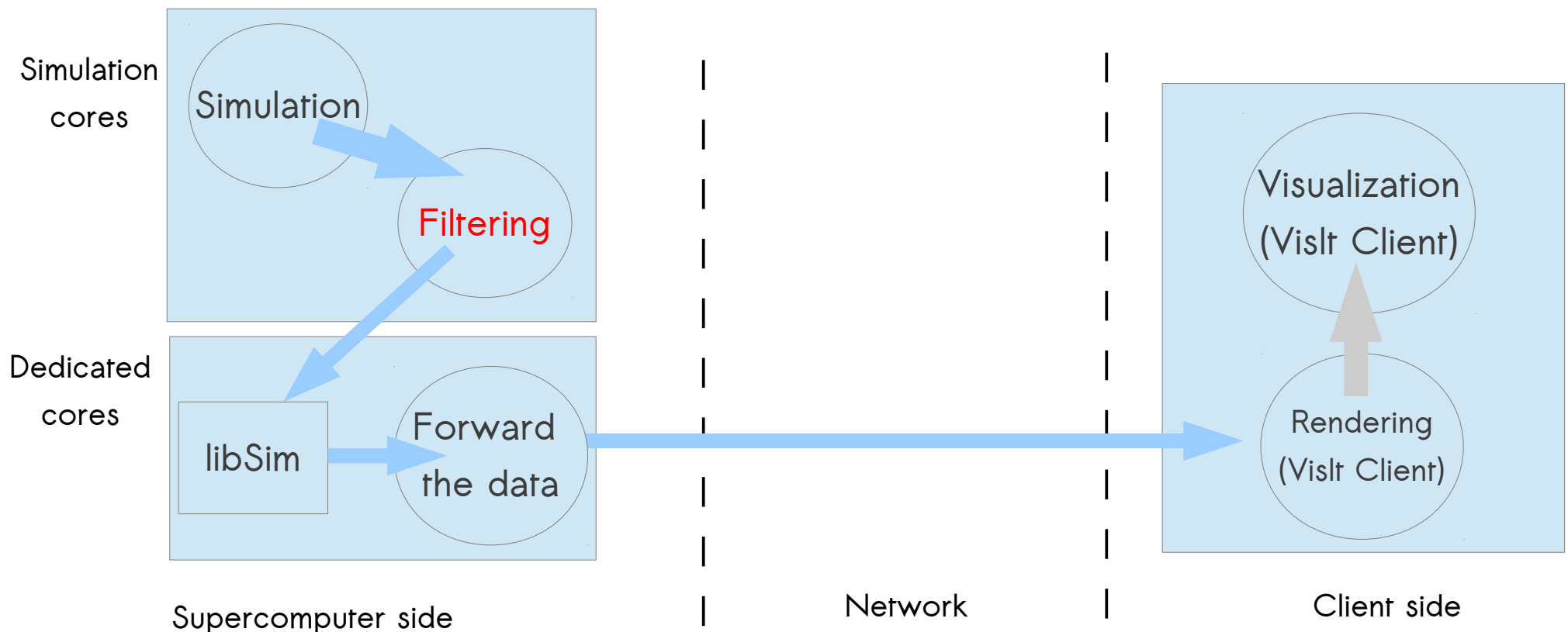
Integration int Existing ISV Frameworks

- Damaris/Viz ISV framework:
 - Damaris : I/O framework
 - VisIt : Client-Server visualization tool



Integration int Existing ISV Frameworks

- Damaris/Viz ISV framework:
 - Damaris : I/O framework
 - VisIt : Client-Server visualization tool

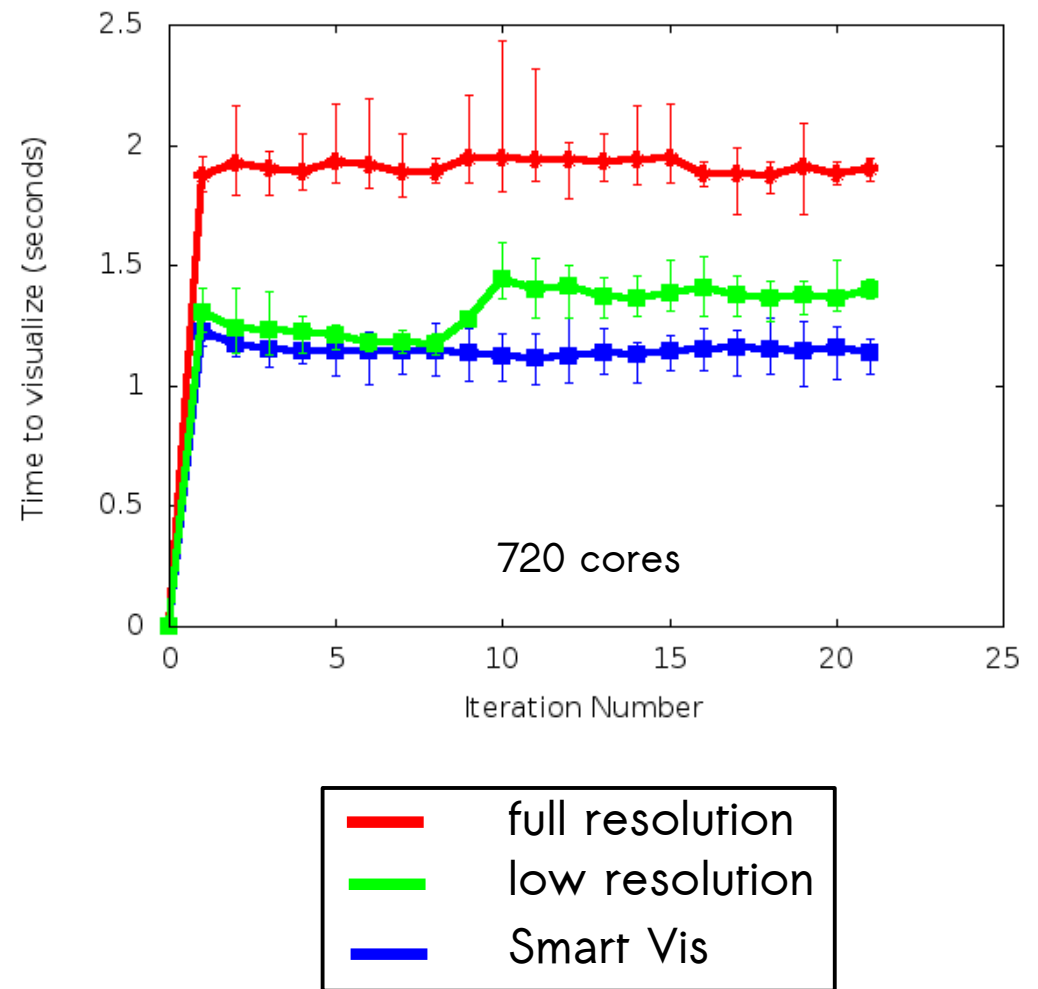


Outline

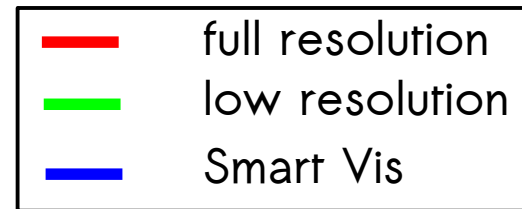
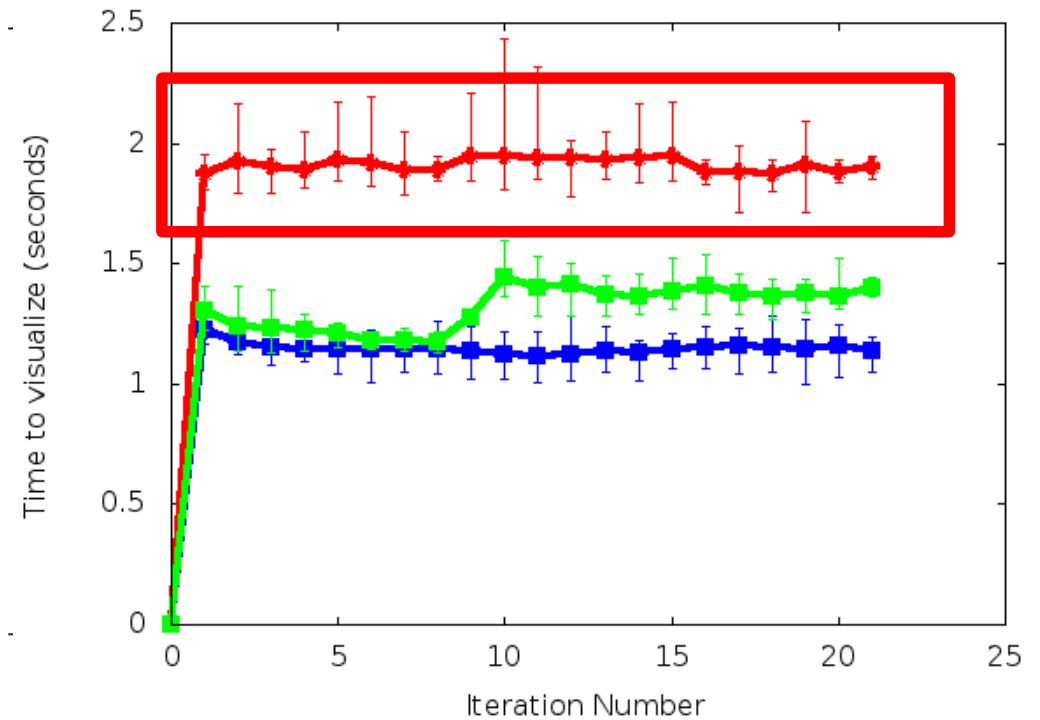
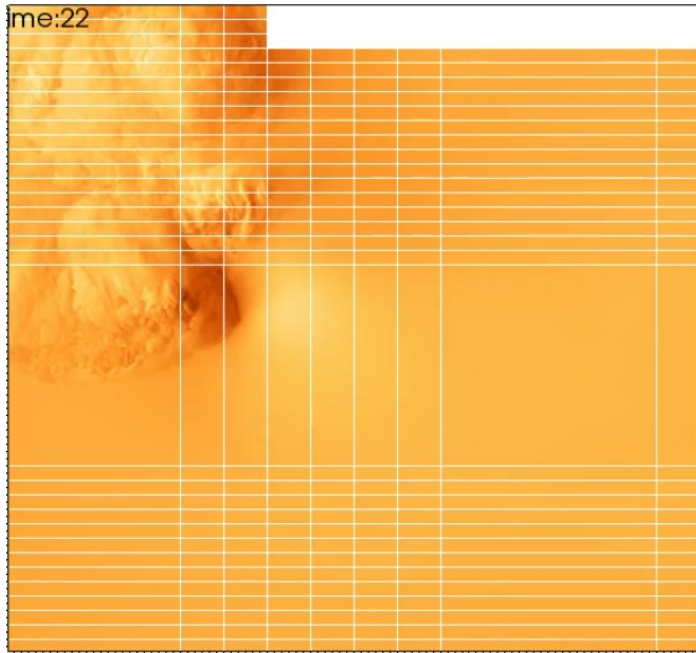
1. HPC Climate Simulations & In Situ Visualization
2. Smart In Situ Visualization
3. Experimental Results
4. Conclusion

CM1 on Grid5000

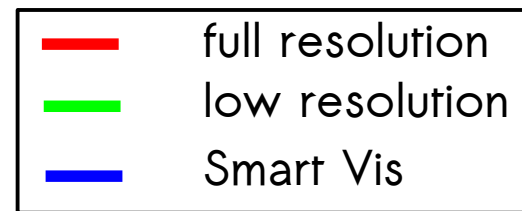
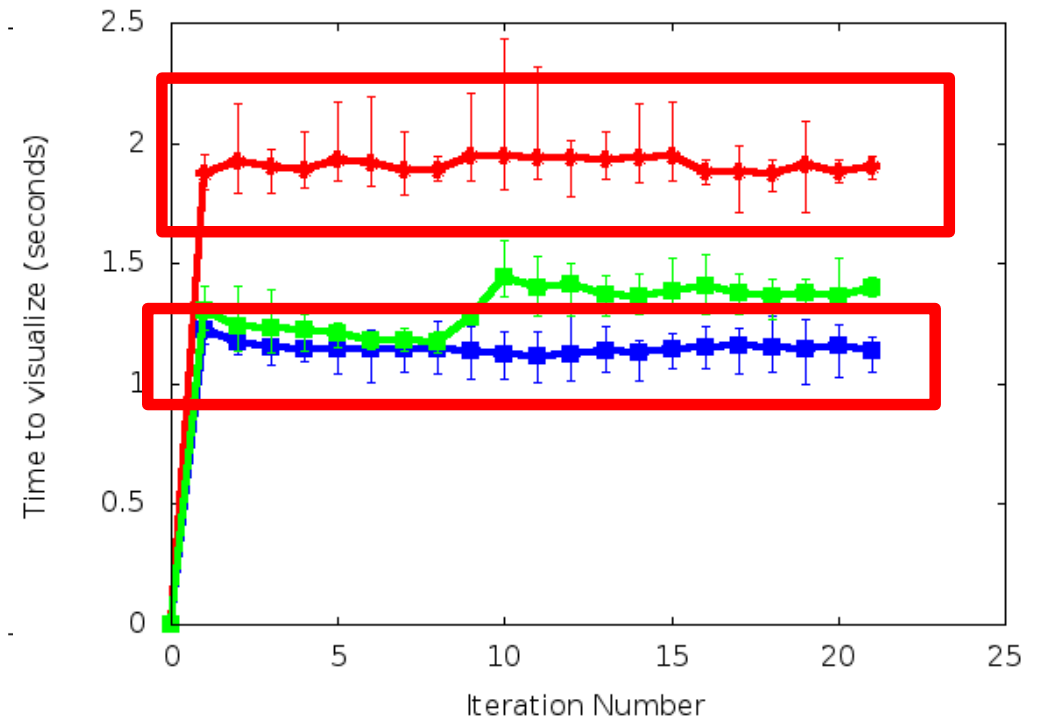
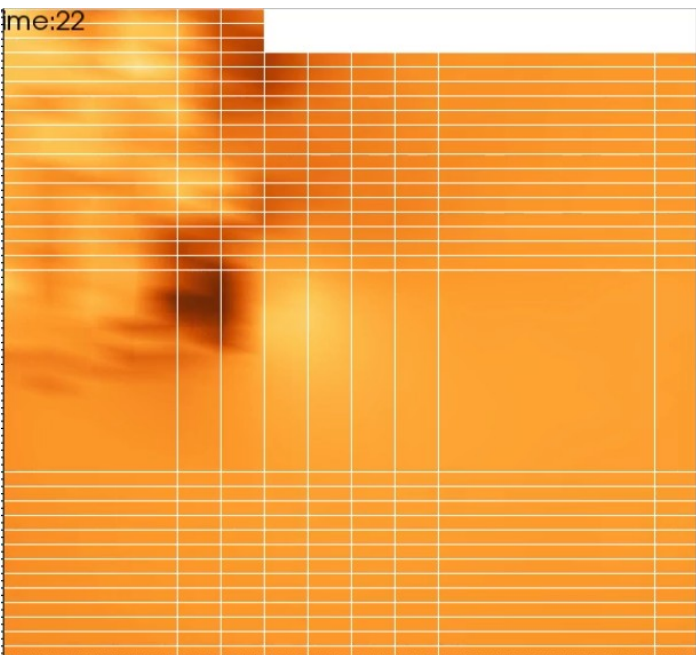
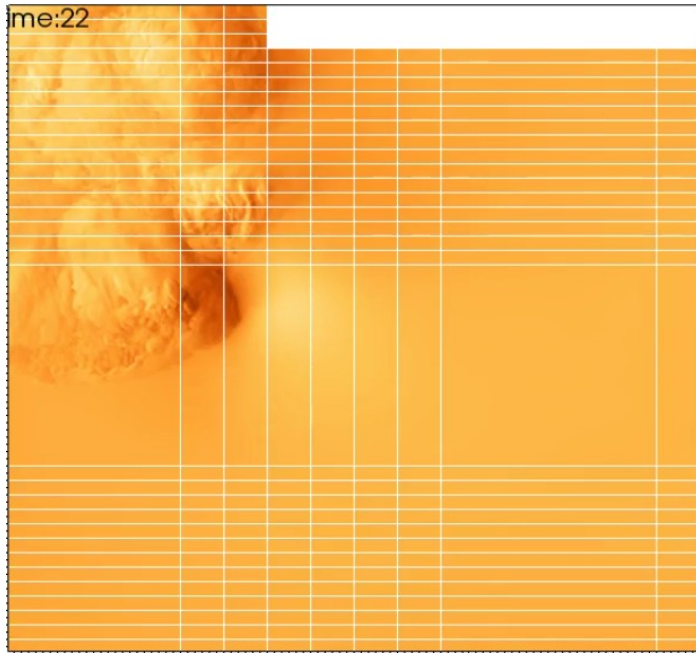
- CM1 on Reims cluster:
 - 30 nodes, 24 cores each
 - 1Gbits/s Ethernet



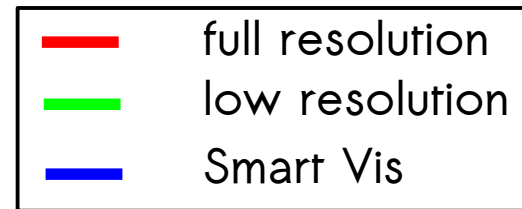
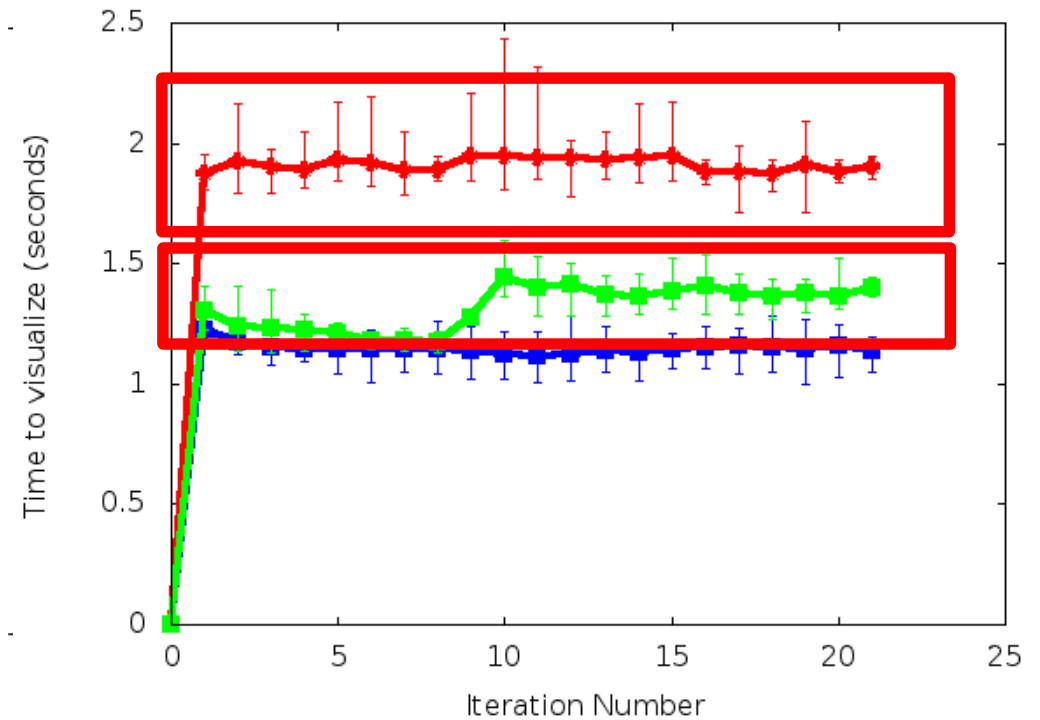
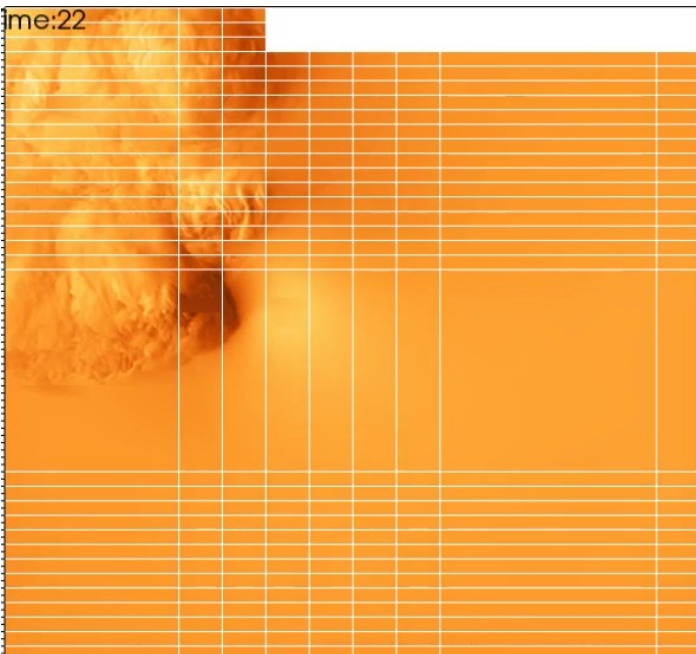
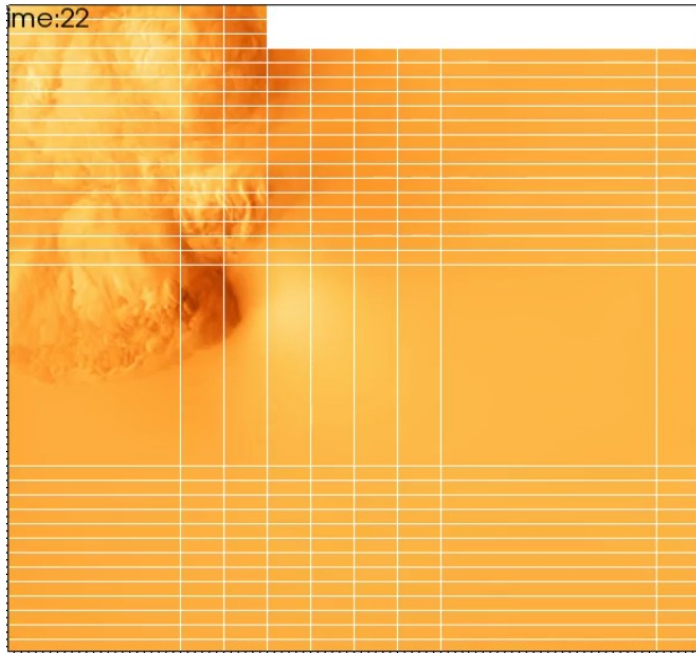
CM1 on Grid5000



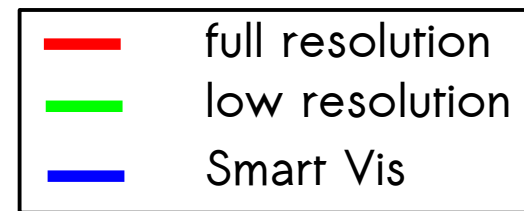
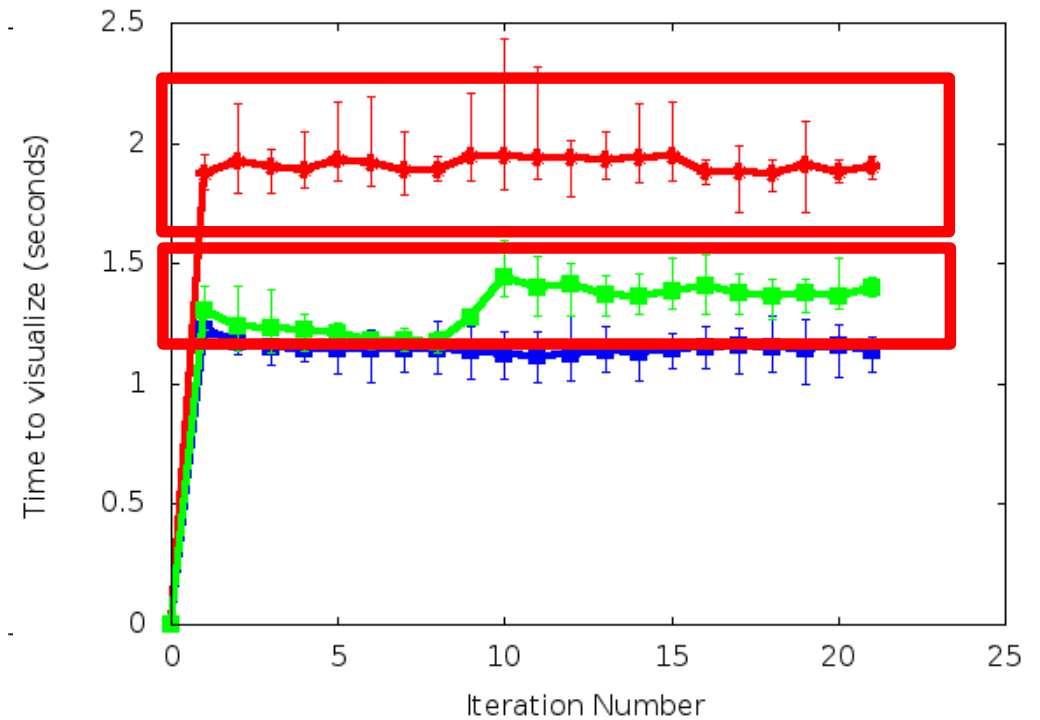
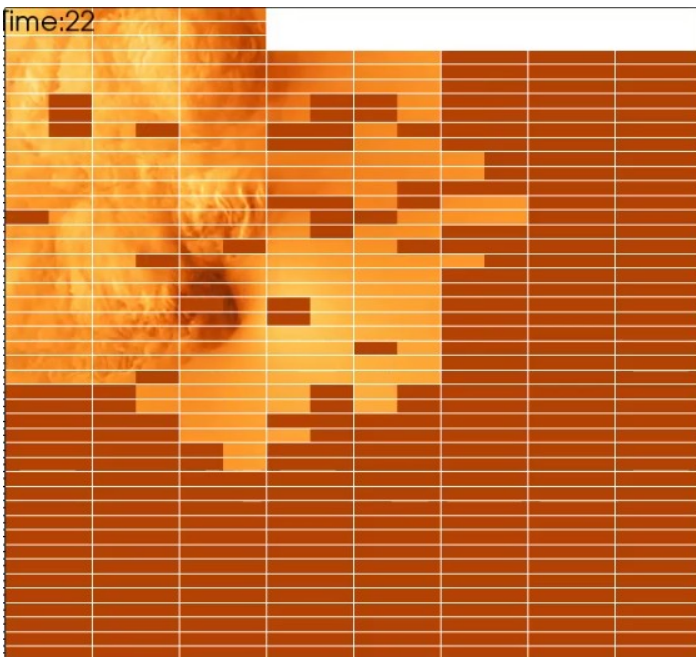
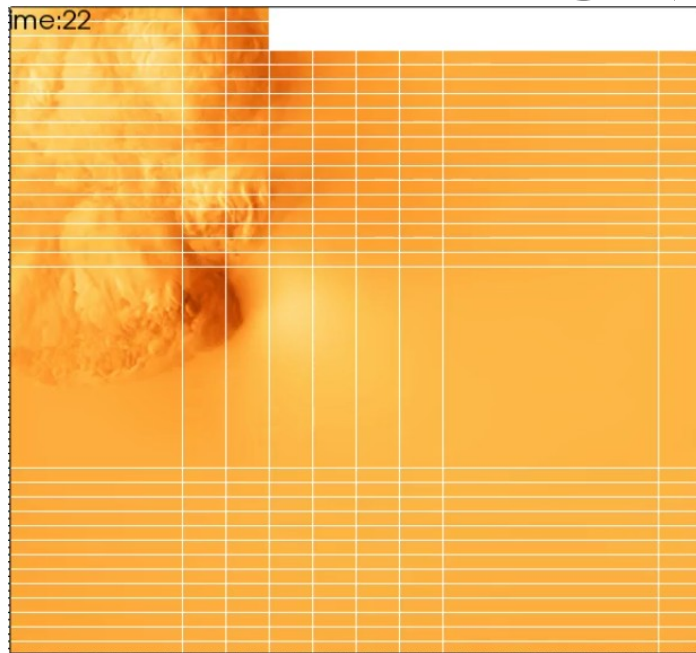
CM1 on Grid5000



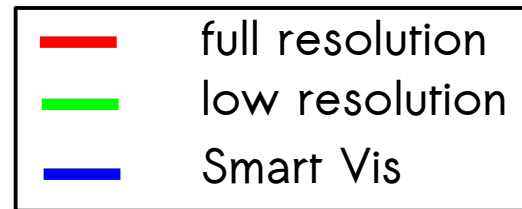
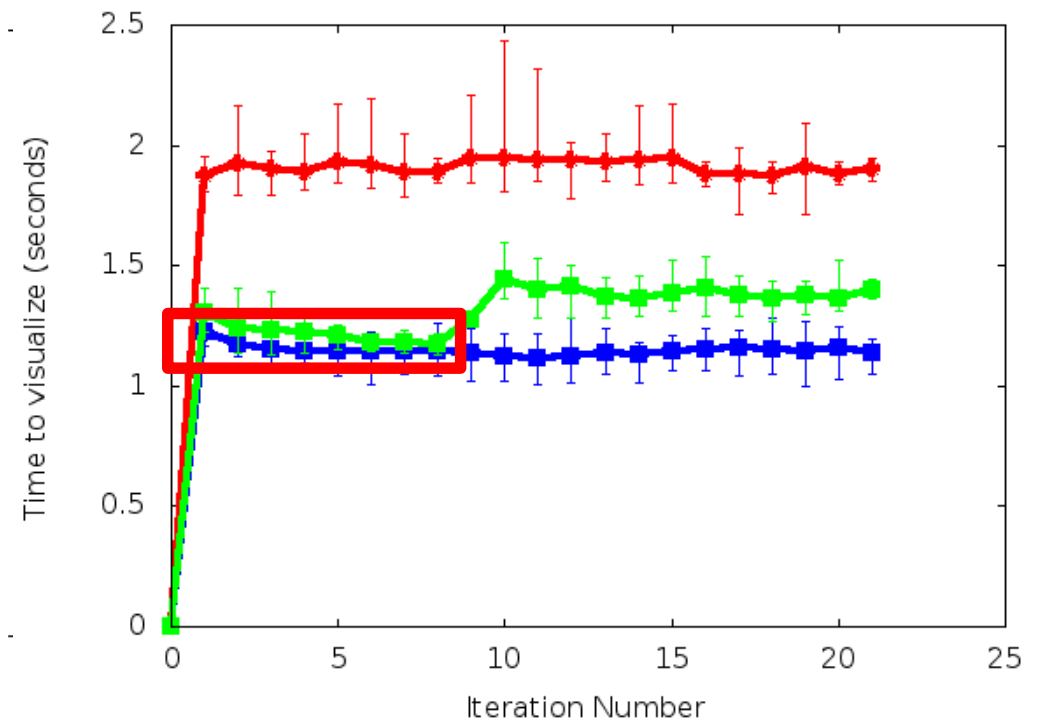
CM1 on Grid5000



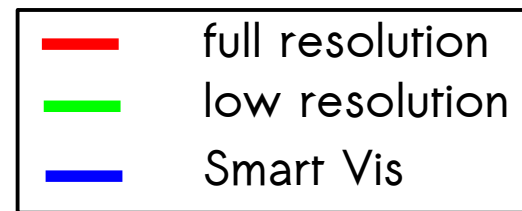
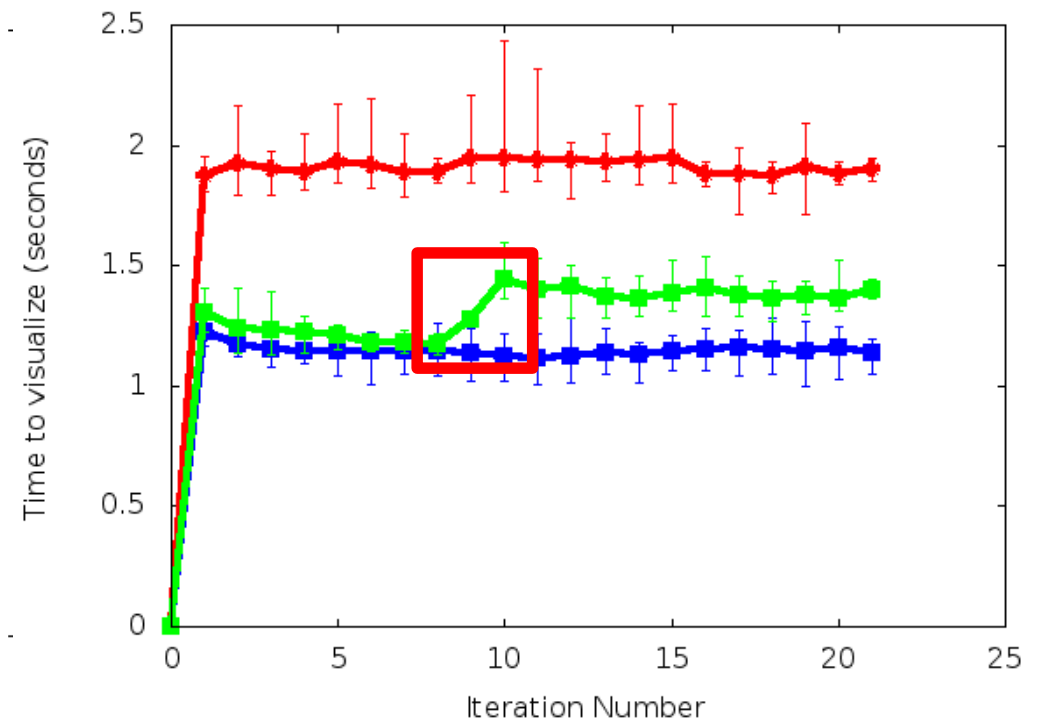
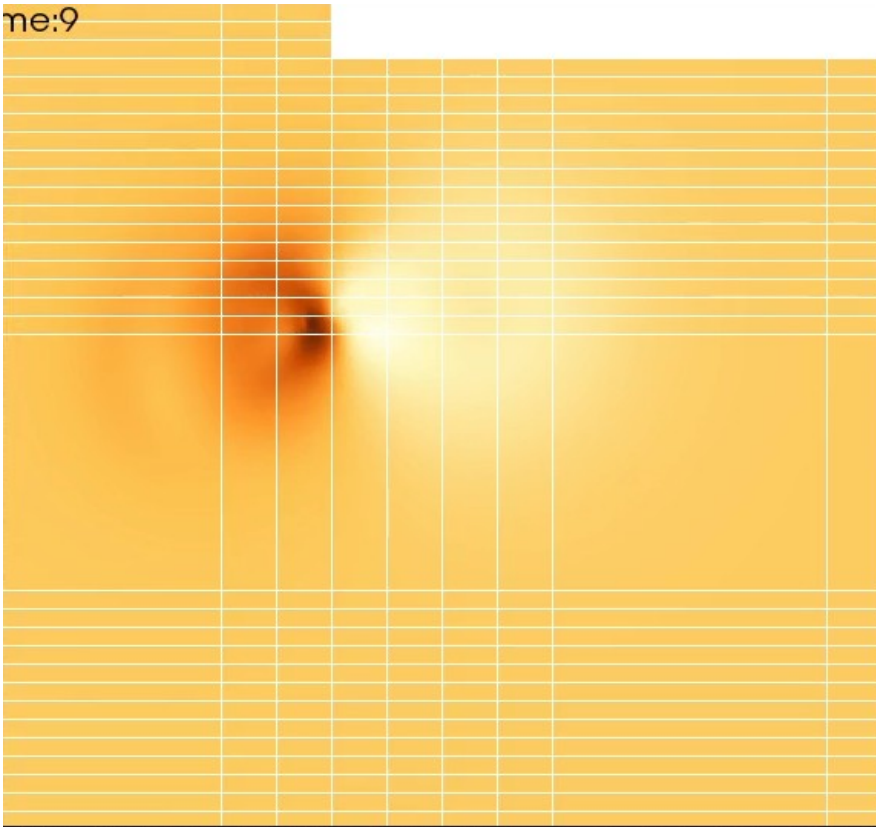
CM1 on Grid5000



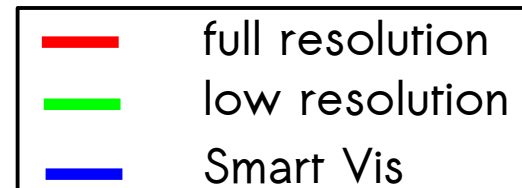
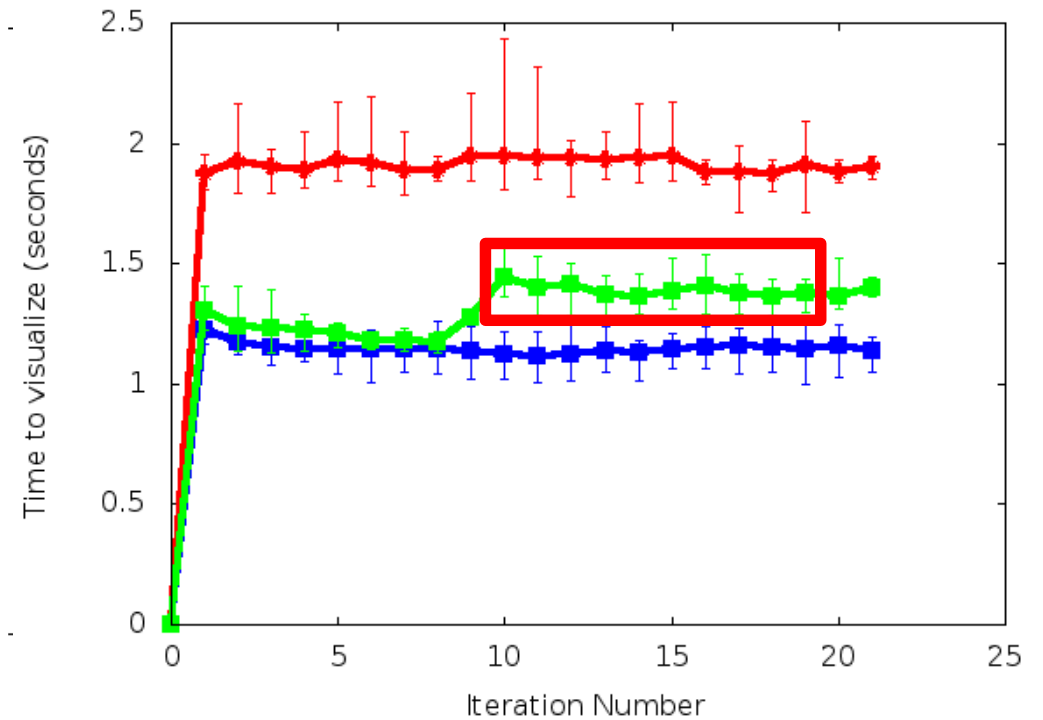
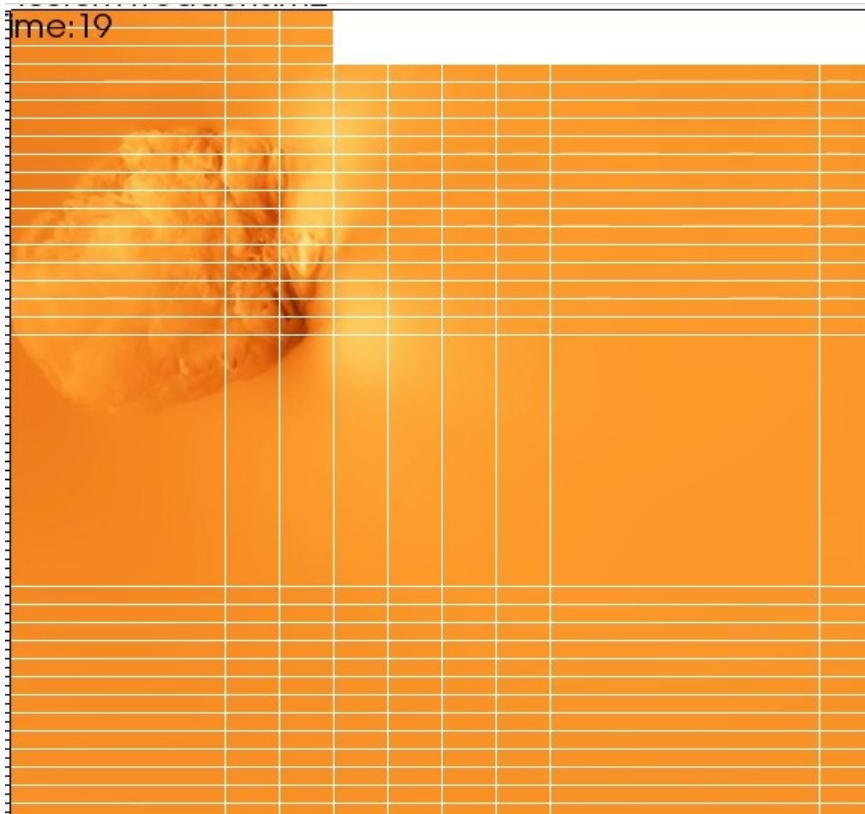
CM1 on Grid5000



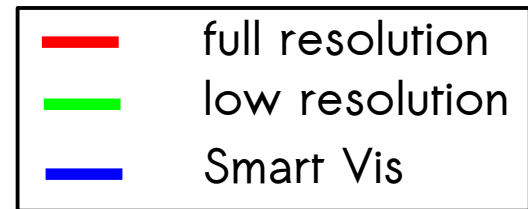
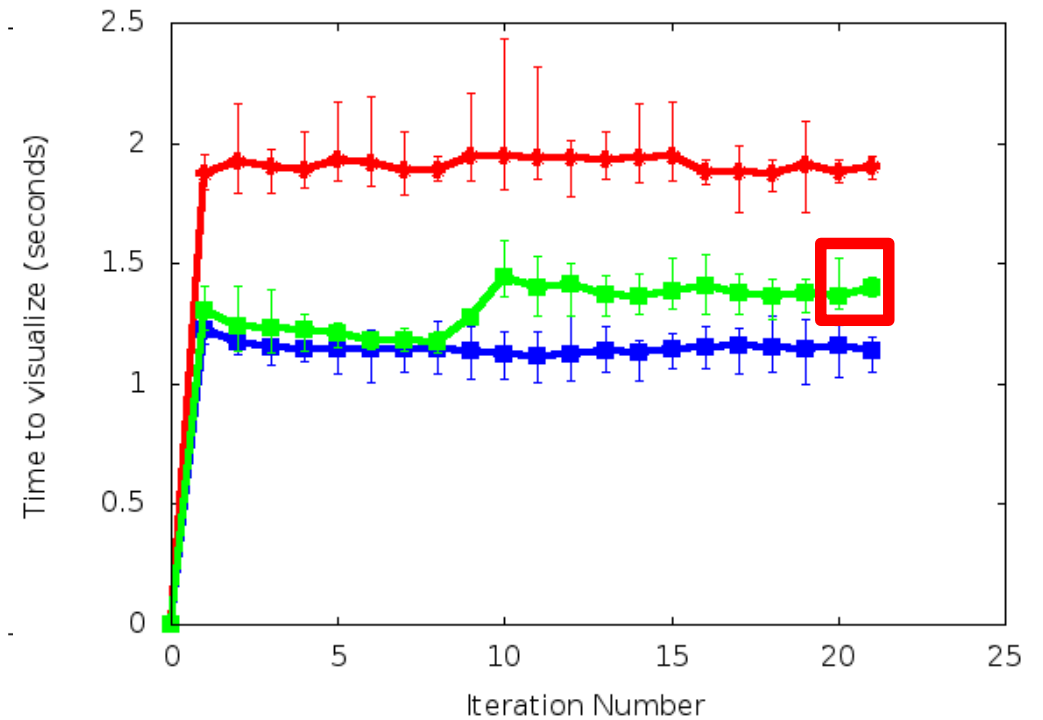
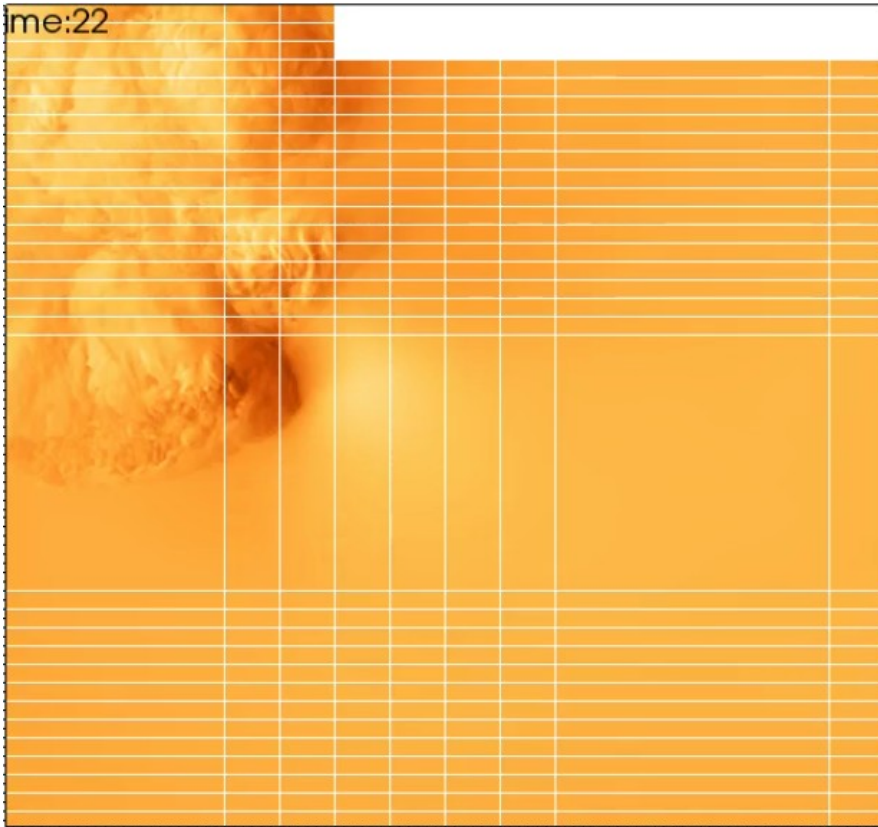
CM1 on Grid5000



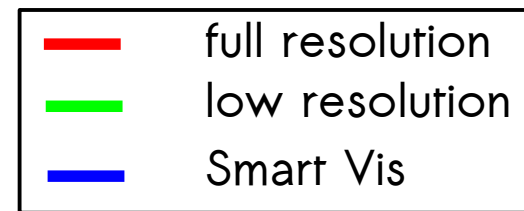
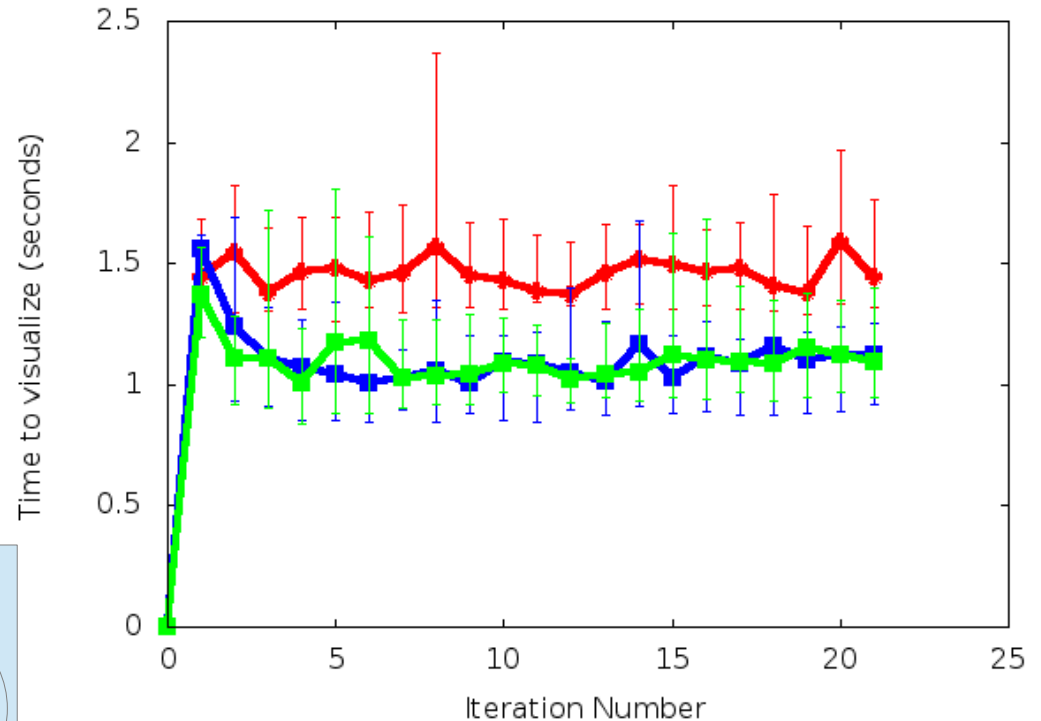
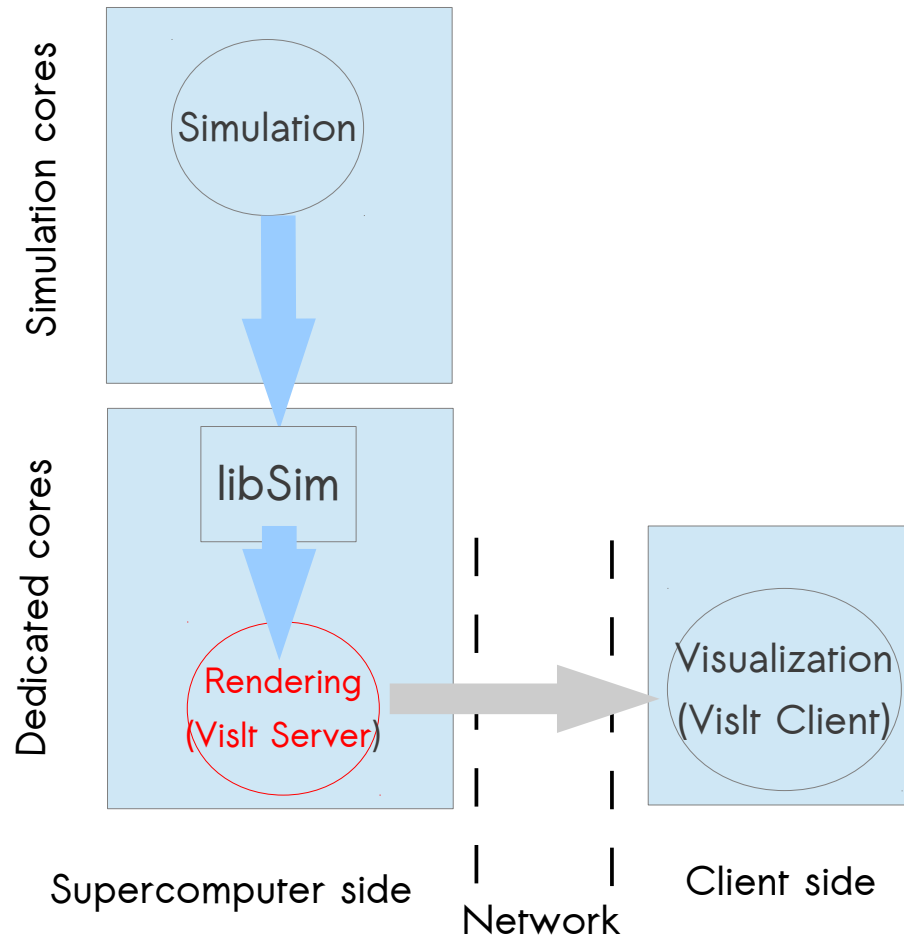
CM1 on Grid5000



CM1 on Grid5000



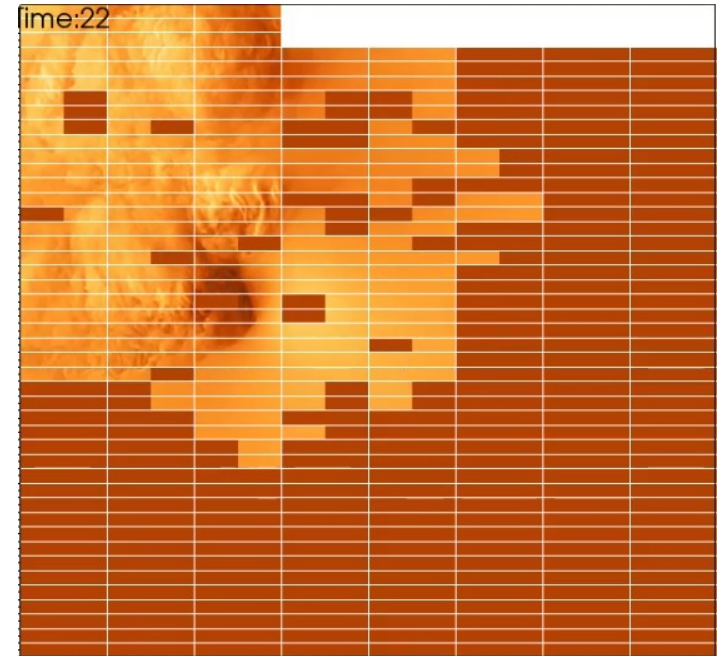
CM1 on Grid5000



VisIt Scalable Rendering

Gaining the Trust of Scientists

- Provide a window to show filtered areas
- Implement a VisIt plug-in with :
 - A slider to control the metrics
 - A list of available metrics
- Provide a metric to calculate the QoV
 - Objective metrics (ex MSRE, MSR) : bad metrics
 - Working with image processing researchers on a best metric



☒ Entropy

☐ Coefficient of

☐ Other

☒ Use gradient

Precision

0100%

Outline

1. HPC Climate Simulations & In Situ Visualization
2. Smart In Situ Visualization
3. Experimental Results
4. Conclusion

Conclusion

- Defined a semantic based on variation in the data
- Propose different metrics to automatically detect relevant data
- Integrate it in existing ISV framework in a complete transparent way
- Provide a gain up to 40% with no considerable loss in quality of visualization

On going work

- Implement a slider to control the metrics as a plug-in in VisIt
- Propose a metric to calculate the quality of visualization QoV => Hard to use an objective metric
- Validation tests at larger scale on BlueWaters are being conducted